

# A Hipótese de Kuznets para os Municípios Brasileiros: Testes para as Formas Funcionais e Estimações Não-Paramétricas

**Erik Alencar de Figueiredo**

*Professor do Departamento de Economia, Programa de Pós-Graduação em Economia, Universidade Federal da Paraíba (UFPB) e Pesquisador do CNPq, Brasil*

**Julio César Araújo da Silva Junior**

*Aluno do Curso de Mestrado em Economia, Universidade Federal do Rio Grande do Sul (UFRGS), Bolsista CNPq, Brasil*

**Paulo Andrade Jacinto**

*Professor do Departamento de Economia, Programa de Pós-Graduação em Economia do Desenvolvimento, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) e Pesquisador do CNPq, Brasil*

---

## Resumo

O objetivo deste estudo é testar a hipótese de Kuznets para os municípios brasileiros nos anos de 1991 e 2000. Para tanto, adotou-se uma estratégia empírica baseada em três etapas: i) utilizam-se os testes para especificações quantílicas desenvolvido por Racine (2006); ii) o método de estimação não-paramétrica *B-splines* proposta por Koenker et alii (1994) e iii) o teste de razão de verossimilhança de Fan et alii (2001). Os resultados apontaram para a inconsistência da estrutura quantílica linear para o conjunto de dados do estudo. As estimativas não-paramétricas sugeriram a validade da hipótese de Kuznets. Por fim, os testes de razão de verossimilhança indicaram a superioridade estatística dos modelos não-paramétricos em relação aos lineares.

*Palavras-chave:* Hipótese de Kuznets, Municípios Brasileiros, Testes para Especificação, *B-Splines*, Razão de Verossimilhança

*Classificação JEL:* C14, C21, O11, O15

---

## Abstract

The main aim of this paper was to analyse the existence of Kuznets hypothesis for the Brazilian Municipalities from 1991 to 2000. The empirical strategy was based on three steps. The first one consists in testing the quantile specification proposed by Racine (2006). The second was to use the non-parametrical estimation *B-splines* according to Koenker et alii (1994). Finally, the third step was to test Fan et alii (2001) Likelihood ratio. The

results show that the quantile linear specification is inconsistent. The non parametrical estimation confirms the Kuznets hypothesis. The Likelihood ratio proves the goodness to fit of the non parametrical estimations comparing to the linear ones.

---

## 1. Introdução

A hipótese da relação não-linear entre os níveis de desigualdade de renda e crescimento econômico postulada por Kuznets (1955) tem sido objeto de intenso debate na literatura especializada. Essa não-linearidade é descrita a partir de uma curva na forma de um U-invertido, indicando que, inicialmente, o padrão de desigualdade no curto prazo aumenta com o crescimento econômico e, no longo prazo, decresce a partir de um *turning point*. As explicações teóricas utilizadas para demonstrar esse padrão são diversas, baseando-se na transição existente entre os setores agrícola e industrial (Robinson 1976); na mudança dos sistemas de financiamento (Greenwood e Jovanovic 1990); e no progresso tecnológico (Galor e Tsiddon 1997).

Os estudos empíricos procuram captar a não-linearidade da curva de Kuznets a partir de uma equação paramétrica sugerida por Ahluwalia (1976), na qual a desigualdade é explicada por um polinômio do segundo grau da renda per capita. Por sua simplicidade, esta forma funcional passou a ser a especificação preferida nos estudos com intuito de validar ou não a existência de uma curva na forma de U-invertido. Entre esses estudos pode-se citar Deininger e Squire (1998), Barro (2000), Thornton (2001), Iradian (2005), entre outros.

Outras especificações podem ser encontradas na literatura sobre o tema. Anand e Kanbur (1993b,a) derivaram uma relação entre desigualdade de renda e crescimento econômico para as seis principais medidas de desigualdade, apresentando as condições necessárias para que a curva no formato de U-invertido fosse obtida. List e Gallet (1999) foram além dessas especificações, buscaram saber o que ocorria após a curva de Kuznets. Para tanto, utilizaram um polinômio de terceiro grau para a renda per capita. Os resultados sugeriram que a partir de um determinado nível de renda per capita, a desigualdade volta a subir.

Em suma, há uma grande variedade de formas funcionais fomentando a discussão referente à validade ou não da hipótese de Kuznets.<sup>1</sup> Sob esse cenário, a econometria não-paramétrica surgiu como um caminho natural, dado que ela é livre de imposições referentes ao processo gerador dos dados. Por isso, diversos estudos utilizaram-se desse instrumental como Huang (2004), Huang et alii (2007) e Lin et alii (2006), entre outros. Contudo, seus resultados apoiaram-se na suposta superioridade dos modelos não-paramétricos, sem a realização de testes que

---

\* Recebido em julho de 2009, aprovado em fevereiro de 2011.

E-mail addresses: eafigueiredo@gmail.com, julio\_economia@yahoo.com.br, pajjap@hotmail.com

<sup>1</sup> Uma lista de estudos empíricos é apresentada em Fields (2001).

confirmem a sua robustez. Ou seja, não são realizados testes para as especificações paramétricas, tampouco há uma comparação estatística entre os instrumentais utilizados. Assim, o presente trabalho tem como objetivo verificar a hipótese do U-invertido entre a desigualdade de renda e crescimento econômico para os municípios do Brasil no período de 1991-2000, fazendo uso de testes para as formas funcionais paramétricas e estimações não-paramétricas.

A despeito dos estudos realizados no Brasil sobre a curva de Kuznets, como o de Lledó (1996) que apresentou uma análise sobre distribuição de renda, crescimento endógeno e política fiscal para os estados brasileiros não encontrando evidências para apoiar a hipótese do U invertido; Bagolin et alii (2002), Jacinto e Tejada (2004), Salvato et alii (2006) e Barros e Gomes (2008) que utilizaram modelos paramétricos e cujos resultados indicaram a existência de curva na forma de U-invertido, a contribuição desse estudo para literatura sobre o tema pode ser vista em duas direções.

A primeira atende a necessidade de buscar mais evidências sobre a curva na forma de U-invertido gerando resultados a partir de regressões quantílicas paramétricas conjuntamente testes para verificar se especificações utilizadas são adequadas. A segunda contribuição está em adotar um instrumental não-paramétrico para verificar a existência da curva Kuznets e da aplicação de testes comparativos entre as inferências paramétricas e não-paramétricas. Vale mencionar que o uso de métodos não-paramétricos possibilita uma maior flexibilidade nas estimações, uma vez que, “permite os dados falarem por si mesmo”, possibilitando, ainda, avaliar a robustez dos resultados obtidos pelos modelos paramétricos. De fato, essa constitui a maior contribuição do presente trabalho.

Assim, além dessa breve introdução, o trabalho foi organizado em mais três seções. Na próxima, serão apresentados os modelos empíricos utilizado para testar a validade da curva de Kuznets. Na Seção três são apresentadas e discutidas as evidências a partir da análise de modelos paramétricos e não-paramétricos. A última seção é reservada às conclusões do trabalho.

## 2. Modelos Empíricos

A estratégia empírica do estudo é constituída de três etapas. Em primeiro lugar serão testadas diversas especificações paramétricas para a curva de Kuznets. Nesta etapa destacam-se os testes *kernel* base desenvolvidos por Racine (2006). O principal objetivo dos testes é mensurar a pertinência de uma caracterização quantílica paramétrica para a hipótese de Kuznets. Caso as especificações paramétricas não sejam estatisticamente significativas, será adotada uma modelagem não-paramétrica para os dados. Esta etapa envolve a modelagem quantílica *B-splines* proposto por Koenker et alii (1994). A modelagem não-paramétrica permitirá que “os dados falem por si”, sem a necessidade de impor formas funcionais a priori. Por fim, um teste de comparação entre os dois métodos (paramétricos e não-paramétricos) atestará a robustez dos resultados.

## 2.1. Testes para especificação

A correta especificação de um modelo paramétrico constitui um importante desafio para os processos de estimação e inferência. Sem dúvida, a perfeita especificação representa o “primeiro melhor” para a estratégia empírica, contudo, nem sempre a teoria aponta o caminho para isso (Li e Racine 2007). Diante dessa impossibilidade, é fundamental a adoção de testes apropriados, dado que os resultados paramétricos mostram-se inconsistentes sob uma má especificação. Ciente deste desafio, este estudo opta pelo teste kernel-based desenvolvido por Racine (2006).

Para entender a estrutura do teste de especificação de Racine (2006), considere a representação quantílica paramétrica:

$$y_i = q(x_i\beta(\tau)) + \eta_i, \quad i = 1, \dots, n$$

Caso a especificação seja a correta, a esperança do erro  $\eta_i$  condicionada a  $x_i$ , em cada quantil  $\tau$  será igual a zero. Dessa forma,

$$\hat{\beta}(\tau) = \arg \min_{\beta \in \mathbb{R}^d} \sum_{i=1}^n \rho_{\tau}(y - q(x_i, \beta(\tau))) \quad (2.2)$$

fornece uma estimativa consistente para  $\beta(\tau)$ . Sendo  $\rho_{\tau}(z) = z[\tau - 1(z < 0)]$  uma função de checagem e 1 corresponde a uma função indicadora [para detalhes, ver Koenker e Bassett (1978)]. Contudo, caso  $q(x_i, \beta(\tau))$  seja mal especificada, (2.2) será inconsistente. Ciente disso considere a correta especificação como a hipótese nula:<sup>2</sup>

$$H_0 : P[F(q(x_i, \beta(\tau)) | x_i) = \tau] = 1, \text{ para algum } \beta(\tau) \in \mathbb{N} \forall i \geq 1$$

onde  $q(\cdot, \cdot)$  é uma função conhecida,  $\beta(\tau)$  é um vetor de parâmetros desconhecidos e  $\mathbb{N}$  é um subconjunto compacto de  $\mathbb{R}^d$ . A hipótese alternativa será

$$H_1 : P[F(q(x_i, \beta(\tau)) | x_i) = \tau] < 1, \text{ para algum } \beta(\tau) \in \mathbb{N} \forall i \geq 1$$

A hipótese nula será verdadeira se  $1(y_i \leq q(x_i, \beta(\tau))) = \tau + \varepsilon_i$  e  $E(\varepsilon_i | x_i) = 0$ . Sendo  $\varepsilon_i = 1(y_i \leq q(x_i, \beta(\tau))) - \tau$  e  $f(x_i)$  a densidade marginal para  $X$ , então  $E(\varepsilon_i | x_i) = F(q(x_i, \beta(\tau)) | x_i) - \tau$ . Portanto,  $I = E\{[E(\varepsilon_i | x_i)]^2 f(x_i)\} \geq 0$  e  $I = 0$  se e somente se  $H_0$  for verdadeira. Assim,  $I$  será um candidato para testar  $H_0$ .

Entretanto, uma observação deve ser feita. Regra geral, os métodos kernel são apropriados para dados contínuos. Contudo, utilizando um método de frequências (Li e Racine 2007, cap. 3), pode-se tratar, em um mesmo conjunto de dados, variáveis discretas e contínuas. Sendo assim, defina  $x_i^d$  como um vetor de variáveis discretas de dimensão  $r \times 1$  e  $x_i^c \in \mathbb{R}^d$  como as demais variáveis contínuas. Dessa forma,  $x_i = (x_i^c, x_i^d)$ .

<sup>2</sup> A hipótese nula baseia-se na definição de teste consistente. Para detalhes, ver Li e Racine (2007).

Defina o  $I$  amostral por

$$I_n = n^{-2} \sum_i \sum_{j \neq i} \hat{\varepsilon}_i \hat{\varepsilon}_j K_{\gamma, ij}$$

onde  $K_{\gamma, ij} = W_{h, ij} L_{\lambda, ij}$ ,  $\gamma = h, \varphi$  são os bandwidths,  $W_{h, ij}$  e  $L_{\lambda, ij}$  são as funções kernel multivariadas para os dados discretos e contínuos, respectivamente,  $\hat{\varepsilon}_i = 1(y_i \leq q(x_i, \hat{\beta}(\tau))) - \tau$ , e  $\hat{\beta}(\tau)$  é um estimador  $\sqrt{n}$ -consistente para  $\beta(\tau)$ .

Racine (2006) recomenda a utilização do método de validação cruzada para a seleção dos bandwidths,<sup>3</sup> substituindo  $(h_1, \dots, h_q, \varphi_1, \dots, \varphi_r)$  por suas estimativas  $(\hat{h}_1, \dots, \hat{h}_q, \hat{\varphi}_1, \dots, \hat{\varphi}_r)$ . Com isso, define-se  $\hat{I}_n$ . Sob algumas suposições, o autor demonstra que, sob  $H_0$ ,

$$\hat{J}_n = n \left( \hat{h}_1, \dots, \hat{h}_q \right)^{1/2} \hat{I}_n / \sqrt{\hat{\Omega}} \rightarrow N(0, 1) \quad (2.3)$$

com

$$\hat{\Omega} = \frac{2(\hat{h}_1, \dots, \hat{h}_q)}{n^2} \sum_i \sum_{j \neq i} \hat{\varepsilon}_i^2 \hat{\varepsilon}_j^2 W_{h, ij}^2 L_{\lambda, ij}^2$$

Os valores críticos são tabulados via *wild bootstrap*. Para detalhes relacionados ao processo de inferência, bem como a performance do teste em amostras finitas são apresentados e discutidos em Racine (2006). Grosso modo, o teste consiste em diversas etapas, destacando-se a estimação de regressões com *mixed data* (Racine e Li 2004), a obtenção de estimativas para os bandwidths de variáveis discretas e contínuas.

## 2.2. Modelo quantílico B-splines e teste de razão de verossimilhança

Caso os testes da subseção anterior constatem erro de especificação do modelo paramétrico, o estudo optará pela modelagem quantílica não-linear conhecida como constrained *B-spline smoothing* (COBS). Esse instrumental é derivado do spline quantílico proposto por Koenker et alii (1994). Em resumo, este método de inferência busca minimizar a função

$$\min_{g \in \zeta} \sum_{i=1}^N \rho_{\tau}(y_i - g(x_i))^2 - \lambda \int |g''(x)| dx \quad (2.4)$$

onde  $\zeta$  corresponde a um Sobolev Space para funções diferenciáveis até a segunda ordem,  $g$  é uma função não-paramétrica,  $\lambda$  é um parâmetro de alisamento e  $g''$  é a segunda derivada de  $g$  em relação a  $x$ . Como em todo processo de inferência não-paramétrica, a escolha do parâmetro de alisamento toma lugar de destaque. Seguindo uma orientação de Koenker et alii (1994), o  $\lambda$  será aquele que otimize um critério de informação, em nosso caso, utilizar-se-á o critério de Akaike. Por fim,

<sup>3</sup> Para detalhes relacionados ao método de validação cruzada, ver Härdle (1990).

a solução para (2.4) é fornecida por He e Ng (1999), por meio de um processo de otimização linear.

Todavia, os resultados obtidos em (2.4) são mais robustos do que os lineares? Uma forma de compará-los é fornecida por Fan et alii (2001) a partir do seguinte teste de razão de verossimilhança generalizada:

$$GLR = \frac{N}{2} \frac{SQR - SQIR}{SQR} \quad (2.5)$$

onde  $SQR$  é a soma dos quadrados dos resíduos do modelo linear e  $SQIR$  é a soma dos quadrados dos resíduos do modelo não-paramétrico. Sob a hipótese nula de igualdade entre os dois métodos, Fan e Yao (2003) obtém a distribuição assintótica para (2.5) a partir do *bootstrap*.

### 3. Desigualdade e Crescimento Econômico: O Que os Dados Têm a Dizer?

Nessa seção são apresentados os resultados da estimação de uma curva de Kuznets para os municípios do Brasil considerando a forma reduzida de Ahluwalia (1976), a especificação recomendada por Anand e Kanbur (1993b) para o índice de desigualdade de Gini e a especificação utilizada por List e Gallet (1999). A metodologia empregada compreende aos modelos paramétricos e os modelos não-paramétricos.

#### 3.1. Base de dados

Imagina-se que a forma mais adequada para verificar a existência da curva de Kuznets para os municípios do Brasil seria por meio de um estudo individual de cada município ao longo do tempo, porém em virtude da inexistência de informações que permitam construir séries longas de índices de desigualdade e renda per capita, optou-se pela comparação em diferentes estágios do desenvolvimento para estudar a sua evolução por meio de um pool de dados. Trata-se do mesmo procedimento observado nos estudos empíricos mais recentes que buscam verificar a validade ou não da hipótese da curva na forma de U invertido.

As informações utilizadas para estimar a curva de Kuznets foram obtidas do Atlas de Desenvolvimento Humano (2003) – PNUD, elaborado pela Fundação João Pinheiro e referem-se aos municípios do Brasil relativo aos anos 1991 e 2000. Na construção da base foram realizados os ajustes para as emancipações municipais que ocorreram ao longo desse período. Os valores monetários da renda per capita encontram-se expressos em termos reais de primeiro de agosto de 2000. O índice de Gini mede o grau de desigualdade de renda e a renda per capita é usada como *proxy* para o nível de desenvolvimento econômico.

### 3.2. Evidências da curva de Kuznets a partir de modelos paramétricos

Para encontrar estimativas para a curva de Kuznets será empregado Mínimos Quadrados Ordinários (MQO) e regressões quantílicas paramétricas em três formas funcionais já abordadas na literatura, são elas:

$$\text{Modelo A} \quad G = \alpha + \beta_1 Y + \beta_2 Y^2 + \varepsilon$$

$$\text{Modelo B} \quad G = \alpha + \beta_1 Y + \beta_2 Y^2 + \beta_3 Y^3 + \varepsilon$$

$$\text{Modelo C} \quad G = \alpha + \beta_1 Y + \beta_2 (1/Y) + \varepsilon$$

onde  $G$  é o índice de Gini,  $Y$  o logaritmo da renda per capita e  $\varepsilon$  o termo estocástico normalmente distribuído com média zero e variância constante. A existência da hipótese de Kuznets depende da significância e do sinal dos parâmetros relacionados às variáveis de renda. Para que o Modelo **A** apresente forma de U-invertido é necessário que os coeficientes possuam sinais de  $\beta_1 > 0$  e um  $\beta_2 < 0$ . Por sua vez, o Modelo **B** valida a hipótese de Kuznets se  $\beta_1 > 0$ ,  $\beta_2 < 0$  e  $\beta_3 > 0$ , enquanto na formulação **C** espera-se um  $\beta_2 < 0$ .

A estratégia empírica adotada seguirá as seguintes etapas:

- 1) estimam-se as equações quantílicas paramétricas;
- 2) realizam-se testes para estas especificações;
- 3) no caso da rejeição da hipótese da correta especificação quantílica paramétrica, adota-se o ferramental não-paramétrico;
- 4) por fim, realizam-se testes comparativos para os dois processos de inferência.

A Tabela 1 apresenta as estimativas para a curva de Kuznets para os Modelos A, B e C obtidas a partir do estimador de média condicional e para as regressões quantílicas paramétricas. Observa-se que nas regressões da média condicional os coeficientes são estatisticamente diferentes de zero ao nível de 1% de significância e apresentam os sinais esperados, sugerindo que as condições necessárias para uma curva na forma de U-invertido são atendidas. Esse resultado não é diferente dos já obtidos na literatura sobre o tema usando dados para o Brasil ou para regiões do país empregando uma ou as três especificações utilizadas nesse estudo, como pode ser visto em Bagolin et alii (2002), Barros e Gomes (2008), Salvato et alii (2006) e Jacinto e Tejada (2004).

Ainda na Tabela 1, são apresentadas as estimativas para os quantis  $\tau = 0.1, 0.25, 0.5, 0.75$  e  $0.90$  obtidas pelas regressões quantílicas paramétricas. Nota-se que todos os coeficientes possuem os sinais esperados e são significativos a 1% com exceção do quantil  $0.90$  do Modelo A. Esses resultados sugerem a existência de uma curva na forma de U-invertido validando a hipótese de Kuznets.

Os resultados, no geral, permitem três observações. A primeira, que a semelhança dos resultados apresentados pelas regressões na média condicional, obtiveram-se evidências favoráveis a hipótese de Kuznets para a maioria das distribuições condicionais do Gini. A segunda, é que a magnitude dos coeficientes é diferente quando comparamos os resultados apresentados pela média condicional e pelas

Tabela 1  
Regressões quantílicas lineares

		MQO		Quantis			
		0.10	0.25	0.50	0.75	0.90	
Modelo A							
Y	0.2702*** (0.0076)	0.3102*** (0.0160)	0.3298*** (0.0177)	0.3011*** (0.0178)	0.2549*** (0.0233)	0.0000 (0.0342)	
Y <sup>2</sup>	-0.0799*** (0.0037)	-0.0320*** (0.0016)	-0.0338*** (0.0018)	-0.0304*** (0.0018)	-0.0256*** (0.0024)	0.0000 (0.0035)	
Modelo B							
Y	0.4663*** (0.0162)	2.6405*** (0.1321)	2.7606*** (0.1317)	2.4806*** (0.1276)	2.3756*** (0.1822)	2.7455*** (0.2778)	
Y <sup>2</sup>	-0.2795*** (0.0151)	-0.5237*** (0.0276)	-0.5459*** (0.0269)	-0.4891*** (0.0263)	-0.4683*** (0.0385)	-0.5573*** (0.0574)	
Y <sup>3</sup>	0.0521*** (0.0038)	0.0342*** (0.0019)	0.0355*** (0.0018)	0.0318*** (0.0017)	0.0304*** (0.0026)	0.0373*** (0.0039)	
Modelo C							
Y	-0.0569*** (0.0038)	-0.1553*** (0.0075)	-0.1606*** (0.0090)	-0.1470*** (0.0084)	-0.1298*** (0.0105)	-0.0368*** (0.0182)	
(1/Y)	-0.0201*** (0.0009)	-3.5111*** (0.1579)	-3.6940*** (0.1945)	-3.4755*** (0.1833)	-3.0988*** (0.2370)	-0.8639*** (0.4219)	

Nota: Erros padrões entre parêntesis.

\*\*\*, \*\*, \*, denotam, respectivamente, significância estatística a 1%, 5% e 10%.

regressões quantílicas, indicando potenciais ganhos associados com a estimação da distribuição condicional da desigualdade em relação à média condicional.

Por fim, a terceira observação está associada ao Modelo B. Esse modelo foi proposto e estimado por List e Gallet (1999) cujo objetivo era saber o que estava acontecendo após o U-invertido. Verificou-se que após um determinado nível de renda per capita a desigualdade voltava a subir. A principal justificativa para esse comportamento eles encontraram em Bishop et alii (1991) que sugerira que esse resultado era devido à transição de uma economia com base manufatureira em direção a uma economia com base no setor de serviços.

O resultado apresentado na Tabela 1, seja para regressões na média ou para as quantílicas paramétricas, levou em conta as condições necessárias para a existência de uma curva na forma de U-invertido. Porém a principal indagação é qual delas é



a especificação mais adequada ou a mais correta? Para responder a essa indagação, temos os resultados apresentados na Tabela 2. Eles foram obtidos a partir do modelo *kernel-based* desenvolvido por Racine (2006), considerando um *bootstrap* com mil replicações e a seleção do *bandwidth* via validação cruzada.

Tabela 2

Teste para especificação quantílica paramétrica

	Quantis				
	0.10	0.25	0.50	0.75	0.90
Modelo A	27.8261*** (0.0000)	54.9315*** (0.0000)	56.4306*** (0.0000)	32.7490*** (0.0000)	16.7644*** (0.0075)
Modelo B	8.5508*** (0.0000)	19.3592*** (0.0000)	6.2043*** (0.0000)	2.1877*** (0.0000)	0.3410 (0.1002)
Modelo C	20.7147*** (0.0000)	40.9804*** (0.0000)	37.6652*** (0.0000)	26.4158*** (0.0000)	10.6537*** (0.0000)

Nota: *p*-values entre parêntesis.

\*\*\*, \*\*, \*, denotam, respectivamente, a rejeição da hipótese nula a 1%, 5% e 10% de significância.

Os testes indicam a rejeição da hipótese nula que corresponde à correta especificação para todos os modelos paramétricos considerados, com um nível de significância de 1%. Uma exceção é registrada no quantil 0.90 do Modelo B. Entretanto, mesmo diante deste resultado isolado, pode-se concluir que a inferência paramétrica não se adéqua para essa amostra finita.

Os resultados apresentados nas Tabelas 1 e 2 permitem traçar algumas considerações. A primeira, relacionada à condição de concavidade, mostra que ela é atendida para uma grande parte dos modelos estimados, sugerindo evidências favoráveis à existência de uma curva de Kuznets, independente se o estimador é uma regressão na média condicional ou se é para regressões quantílicas. Exceto para o quantil 0.90 do Modelo A.

Uma segunda consideração, relacionada ao uso de regressões na média condicional e regressões quantílicas sugere que embora a primeira seja muito empregada, a segunda mostra-se bastante útil para enriquecer as análises e para suprir deficiências das estimações por MQO,<sup>4</sup> permitindo verificar como as covariadas afetam mais profundamente a desigualdade de renda no período estudado.

Por fim, a terceira consideração está relacionada a teste realizado para verificar a especificação mais adequada, Modelo A, B ou C. Os testes apontaram a rejeição da hipótese nula de correta especificação. Com exceção para o quantil 0.90. Nesse

<sup>4</sup> Além de lidar melhor com o problema de heteroscedasticidade e com a presença de observações discrepantes, a regressão quantílica permite estimar diferentes efeitos das variáveis covariadas na variável dependente.

sentido, por esse critério, teste de Racine (2006), pode-se dizer que a inferência paramétrica não é mais adequada para esta amostra finita. Nesse sentido, vamos usar os dados e testar para uma especificação não-paramétrica.

### 3.3. Evidências da curva de Kuznets a partir de modelos não-paramétricos

A maior parte da teoria econômica não identifica uma forma funcional específica para a relação entre a variável dependente e a variável explicativa numa regressão. Com intuito de evitar má especificação da estrutura paramétrica, iremos considerar uma abordagem alternativa de estimação não-paramétrica. Por isso, nesta subseção serão explorados os resultados de um modelo não-paramétrico. A vantagem desse método reside no fato de que ele “permite os dados falarem por si mesmo”. Ou seja, o procedimento de regressão não-paramétrica poderia ser considerado como se fosse um estágio final de uma análise de dados ou um degrau no caminho confirmatório do processo de Kuznets. Soma-se a isso, não adequação da estrutura paramétrica para o conjunto de dados do estudo.

A partir das considerações acima, propõe-se a utilização de um instrumental não-paramétrico. Optou-se pelo método *B-splines* proposto por Koenker et alii (1994). Dada a sensibilidade da curva *spline* frente ao parâmetro de alisamento ( $\lambda$ ), foram comparadas as estimativas da curva mediana ( $\tau = 0,50$ ), considerando diferentes valores para  $\lambda$ . As estimativas, sumarizadas na Figura A1, em anexo, não apresentam diferenças expressivas entre as curvas. Mesmo assim, adota-se o critério de Akaike como uma forma de selecionar parâmetro de penalização ótimo.

As estimativas para os dados brasileiros ( $\tau = 0,1, 0,25, 0,5, 0,75$  e  $0,90$ ), são resumidas na Figura 1. Observa-se que a relação entre a desigualdade de renda e a renda per capita aponta para a existência de uma curva de Kuznets. Como não podemos realizar observações sobre os resultados no final da curva, por poder existir uma variabilidade das estimativas não-paramétricas, torna-se difícil fazer comparações a estudos que mencionam a volta do aumento da desigualdade depois de verificada uma diminuição, como exemplo o estudo de List e Gallet (1999).

Contudo, o resultado descrito na Figura 1 é mais robusto do que as especificações paramétricas? Com intuito de comparar as inferências paramétricas (três modelos) e não-paramétrica será realizado um teste da razão de verossimilhança generalizada que foi desenvolvido por Fan et alii (2001). A Tabela 3 apresenta o *p-value* para o teste considerando um *bootstrap* com 1000 replicações. Vale lembrar que a hipótese nula a ser testada é de igualdade entre os dois modelos contra hipótese alternativa de que os métodos não-paramétricos são superiores. Nota-se que o modelo não-paramétrico é superior em todos os modelos estimados, dado que a hipótese nula do teste é rejeitada a um nível de significância de 1%.

Esse resultado se deve em grande parte pela flexibilidade do modelo não-paramétrico em que permite o pesquisador obter informações acerca de qual a relação existente entre a desigualdade do Gini e a renda per capita, podendo, inclusive, ser útil como um procedimento para validar a especificação paramétrica

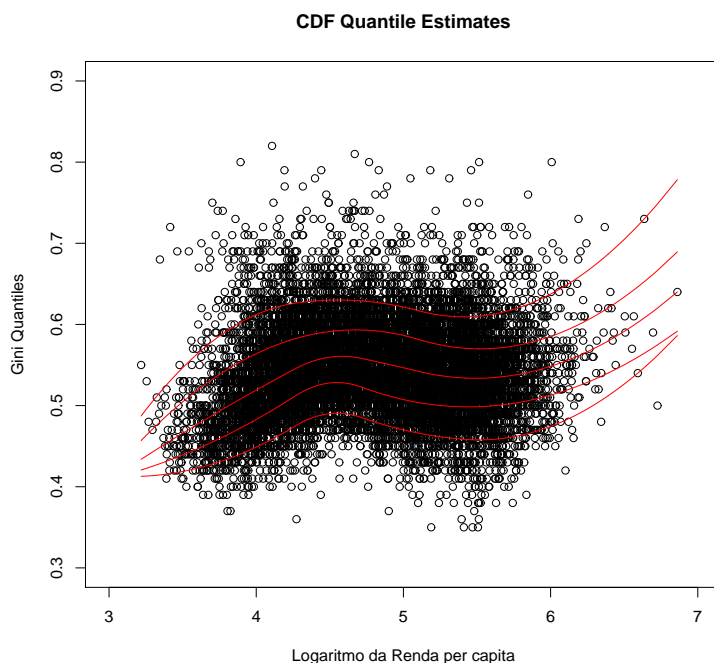


Fig. 1. Estimativas não-paramétricas – Brasil

mais adequada. No presente estudo pelo resultado encontrado mostra que se o pesquisador fizer a opção pelo uso de modelos paramétricos para testar a existência da curva de Kuznets para o Brasil ele não pode deixar de considerar a sua especificação do modelo com o polinômio de terceiro grau como sendo o mais adequado.

Tabela 3

*P*-value do teste de razão de verossimilhança generalizada

Quantis					
Modelo	0.10	0.25	0.50	0.75	0.90
Modelo A	0.0000	0.0000	0.0000	0.0000	0.0000
Modelo B	0.0000	0.0000	0.0000	0.0000	0.0000
Modelo C	0.0000	0.0000	0.0000	0.0000	0.0000

Fonte: Cálculo do autor.

Em resumo, os resultados apresentados nessa seção sugerem que a inferência linear não é consistente e que o modelo não-paramétrico indica a validade da

hipótese de Kuznets. Além disso, mostra que a inferência não-paramétrica é mais apropriada quando comparada aos três modelos paramétricos estimados.

#### 4. Considerações Finais

As controvérsias em torno da validade da hipótese do U invertido foram e tem sido gerada por uma série de estudos na qual dada a natureza das informações, diferentes especificações para a forma funcional e o método econométrico utilizado permitem validar ou não as conclusões de Kuznets. Independente dos resultados, esses estudos cumprem o papel de atender a necessidade de um conhecimento mais sólido e uma perspectiva mais convincente para a economia.

Esse estudo propôs o re-exame da hipótese de Kuznets para os municípios do Brasil dispostos em um pool de dados. A estratégia empírica contou com uma série de instrumentais não-paramétricos, visando testar as formas funcionais lineares, promover uma inferência não-paramétrica, baseadas em *B-splines* e comparar os métodos quantílicos lineares e não-paramétricos.

Os resultados apontaram para a inconsistência da estrutura linear paramétrica para o conjunto de dados do estudo. Os *B-splines* sugeriram a validade da hipótese de Kuznets. Por fim, os testes de razão de verossimilhança indicaram a superioridade estatística dos modelos não-paramétricos em relação aos lineares. Porém, uma ressalva deve ser feita: a estimação não-paramétrica não considerou uma estrutura multivariada. Pois ela necessitaria de estimativas não-paramétricas com *mixed data*. Nesse sentido, o próximo passo consistirá na aplicação do ferramental desenvolvido por Li e Racine (2008).

#### Referências bibliográficas

- Ahluwalia, M. (1976). Income distribution and development: Some stylized facts. *American Economic Review*, 66:128–153.
- Anand, S. & Kanbur, S. (1993a). Inequality and development: A critique. *Journal of Development Economics*, 41:19–43.
- Anand, S. & Kanbur, S. (1993b). The Kuznets process and the inequality-development relationship. *Journal of Development Economics*, 40:25–52.
- Bagolin, I., Gabe, J., & Ribeiro, E. P. (2002). Crescimento e desigualdade no Rio Grande do Sul: Uma revisão da Curva de Kuznets para os municípios gaúchos (1970-1991). In *Anais do XXX Encontro Nacional de Economia*, Nova Friburgo. ANPEC.
- Barro, R. (2000). Inequality and growth in a panel of countries. *Journal of Economic Growth*, 5:5–32.
- Barros, L. & Gomes, F. (2008). Desigualdade e desenvolvimento: A hipótese de Kuznets é válida para os municípios brasileiros? *Revista Análise Econômica*, 50:57–81.
- Bishop, J., Formby, J., & Thistle, P. (1991). Changes in US earnings distributions in the 1980s. *Journal of Applied Economics*, 23:425–34.
- Deiningner, K. & Squire, L. (1998). A new data set measuring income inequality. *World Bank Economic Review*, 10:565–92.

- Fan, J. & Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer.
- Fan, J., Zhang, C., & Zhang, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *The Annals of Statistics*, 29:153–193.
- Fields, G. (2001). *Distribution and Development: A New Look at the Developing World*. MIT Press, Cambridge.
- Galor, O. & Tsiddon, D. (1997). Technological progress, mobility and economic growth. *American Economic Review*, 87:363–382.
- Greenwood, J. & Jovanovic, B. (1990). Financial development, growth and the distribution of income. *Journal of Political Economy*, 98:1076–1107.
- Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press.
- He, X. & Ng, P. (1999). COBS: Qualitatively constrained smoothing via linear programming. *Computational Statistics*, 14:315–337.
- Huang, H. (2004). A flexible nonlinear inference to the Kuznets hypothesis. *Economics Letters*, 84:289–296.
- Huang, H., Lin, S., Suen, Y., & Yeh, C. (2007). A quantile inference of the Kuznets hypothesis. *Economic Modelling*, 24:559–570.
- Iradian, G. (2005). Inequality, poverty and growth: Cross-country evidence. IMF Working Paper.
- Jacinto, P. & Tejada, C. (2004). Desigualdade de renda e crescimento econômico nos municípios da região nordeste do Brasil: O que os dados têm a dizer? In *Anais do XXXII Encontro de Economia*, João Pessoa, PB. ANPEC.
- Koenker, R. & Bassett, G. (1978). Regression quantiles. *Econometrica*, 46:33–50.
- Koenker, R., Ng, P., & Portnoy, S. (1994). Quantile smoothing splines. *Biometrika*, 81:673–680.
- Kuznets, S. (1955). Economic growth and income inequality. *American Economic Review*, 45:1–28.
- Li, Q. & Racine, J. (2007). *Nonparametric Econometrics: Theory and Practice*. Princeton University Press.
- Li, Q. & Racine, J. (2008). Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data. *Journal of Business and Economics Statistics*, 26:423–434.
- Lin, S., Huang, H., & Weng, H. (2006). A semiparametric partially linear inference of the Kuznets hypothesis. *Journal of Comparative Economics*, 34:634–647.
- List, J. & Gallet, C. (1999). The Kuznets' curve: What happens after the inverted-U? *Review of Development Economics*, 3:200–206.
- Lledó, D. (1996). Distribuição de renda, crescimento endógeno e política fiscal: Uma análise cross-section para os estados brasileiros. Texto para Discussão 441. IPEA, Rio de Janeiro.
- Racine, J. (2006). Consistent specification testing of heteroskedastic parametric regression quantile models with mixed data. Working Paper.
- Racine, J. & Li, Q. (2004). Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics*, 119:99–130.
- Robinson, S. (1976). A note on the U-hypothesis relating income inequality and economic development. *American Economic Review*, 66:437–440.
- Salvato, M., Alvarenga, P., França, C., & Araujo Junior, A. (2006). Crescimento e desigualdade: Evidências da Curva de Kuznets para os municípios de Minas Gerais

– 1991/2000. *Economia e Gestão*, 6:1–16.

Thornton, J. (2001). The Kuznets inverted-U hypothesis: Panel data evidence from 96 countries. *Applied Economics Letters*, 8:15–16.

## Anexo A

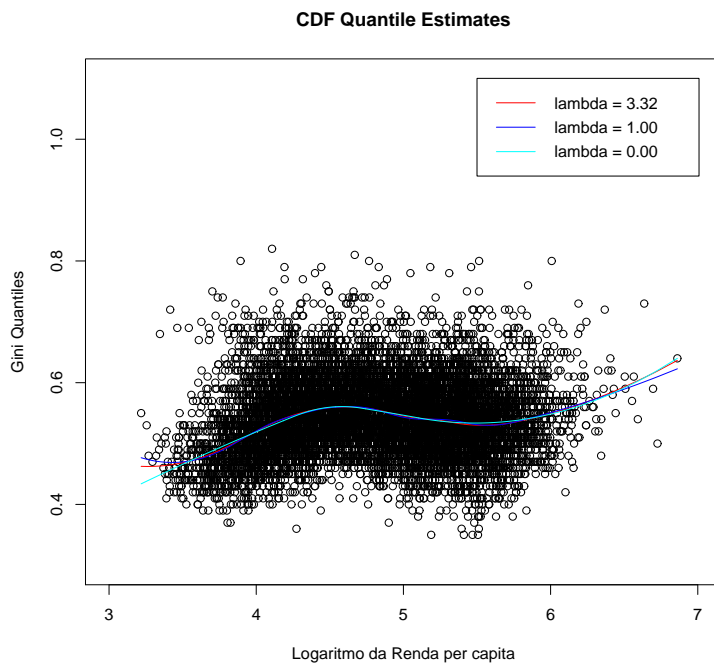


Figura A1: Regressão mediana e parâmetros de penalidade. Lambda ótimo: 3.32